GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: March 9, 2002, 00:09:23 ; Search time 8498.8 Seconds
(without alignments)
29.081 Million cell updates/sec

Title: US-09-851-670-14
Perfect score: 23
Sequence: 1 gagaacacccgctctctcgcaaa 23

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 11351937 seqs, 5372889281 residues

Total number of hits satisfying chosen parameters: 111874

Minimum DB seq length: 0
Maximum DB seq length: 60

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estfun:*
2: em_esthum:*
3: em_estin:*
4: em_estom:*
5: em_estpl:*
6: em_estba:*
7: em_estro:*
8: em_estov:*
9: em_htc:*
10: gb_est1:*
11: gb_est2:*
12: gb_htc:*
13: gb_gss:*
14: em_gss_fun:*
15: em_gss_hum:*
16: em_gss_inv:*
17: em_gss_pln:*
18: em_gss_pro:*
19: em_gss_rod:*
20: em_gss_vrt:*
21: em_gss_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	15.8	68.7	52	10	AW693240	AW693240 NF062A03S
	2	14.6	63.5	50	10	AU107235	AU107235 AU107235
	3	14	60.9	40	10	AI859778	AI859778 wm21e09.x
	4	13.6	59.1	40	13	AZ463268	AZ463268 1M0272B01
c	5	13.4	58.3	34	13	AQ025206	AQ025206 EP(3)1249
	6	13.4	58.3	49	13	AZ514447	AZ514447 1M0361M14
c	7	13.2	57.4	60	13	AZ494986	AZ494986 1M0330A09
c	8	13	56.5	25	13	AZ303786	AZ303786 1M0003G03
c	9	13	56.5	59	10	AA098077	AA098077 mk18g04.r
c	10	12.8	55.7	55	10	AA894695	AA894695 oj27d05.s
	11	12.6	54.8	26	13	AZ342914	AZ342914 1M0076C22
c	12	12.6	54.8	26	13	TA194F01Q	AL477302 T. brucei
c	13	12.6	54.8	35	13	AZ845779	AZ845779 2M0145B13
c	14	12.6	54.8	50	10	AU106768	AU106768 AU106768
	15	12.4	53.9	32	13	AZ462085	AZ462085 1M0269P08
c	16	12.4	53.9	45	13	AZ635879	AZ635879 1M0493O22
	17	12.4	53.9	46	10	AI811483	AI811483 tw43d02.x
c	18	12.4	53.9	49	10	AW432778	AW432778 sh82h02.y
c	19	12.4	53.9	50	10	AW424126	AW424126 sh61c07.y
	20	12.4	53.9	57	13	AQ024986	AQ024986 EP(2)0892
	21	12.4	53.9	58	10	AI499235	AI499235 to08g07.x
	22	12.2	53.0	21	13	AZ581103	AZ581103 1M0369E22
	23	12.2	53.0	21	13	AZ857747	AZ857747 2M0162J08
c	24	12.2	53.0	36	13	BH011404	BH011404 BG01613-5
	25	12.2	53.0	40	13	TA253H01Q	AL483109 T. brucei
c	26	12.2	53.0	49	10	AI792794	AI792794 om81c01.y
	27	12.2	53.0	49	10	AA429584	AA429584 zw77h05.r
	28	12.2	53.0	50	10	AU103878	AU103878 AU103878
	29	12.2	53.0	50	10	AU103887	AU103887 AU103887
c	30	12.2	53.0	50	10	AU106770	AU106770 AU106770
c	31	12.2	53.0	52	11	BF643352	BF643352 NF004A05E
c	32	12.2	53.0	58	13	AZ920130	AZ920130 1006018C0
	33	12	52.2	30	13	AQ025600	AQ025600 l(2)02839
	34	12	52.2	41	11	H93873	H93873 yv08e06.r1
	35	12	52.2	43	10	AI245483	AI245483 qk30b11.x
	36	12	52.2	46	10	AI744926	AI744926 tr17c11.x
c	37	12	52.2	46	13	AZ810625	AZ810625 2M0076H14
	38	12	52.2	49	10	AA663894	AA663894 ae74c09.s
c	39	12	52.2	50	13	AQ025192	AQ025192 EP(3)1161
	40	12	52.2	51	13	AQ025045	AQ025045 EP(2)1176
c	41	12	52.2	51	13	AZ575776	AZ575776 AST-T31B0
c	42	12	52.2	52	10	BE321098	BE321098 NF021G04I
	43	12	52.2	52	11	BF631938	BF631938 NF017G09D
	44	11.8	51.3	26	13	TA65F02Q	AL458692 T. brucei
c	45	11.8	51.3	31	10	AI383811	AI383811 tc98f11.x

ALIGNMENTS

RESULT 1
AW693240
LOCUS AW693240 52 bp mRNA EST 20-DEC-2000
DEFINITION NF062A03ST1F1000 Developing stem Medicago truncatula cDNA clone NF062A03ST 5', mRNA sequence.
ACCESSION AW693240
VERSION AW693240.2 GI:11933549
KEYWORDS EST.
SOURCE barrel medic.
ORGANISM Medicago truncatula
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae; Medicago.
REFERENCE 1 (bases 1 to 52)
AUTHORS He,X.-Z., Shadle,G., Scott,A.D., Harris,A.R., Gonzales,R.A., Bell ,C.J., Flores,H.R., Inman,J.T., Weller,J.W., May,G.D. and Dixon ,R.A.
TITLE Expressed Sequence Tags from the Samuel Roberts Noble Foundation Medicago truncatula stem library
JOURNAL Unpublished (2000)
COMMENT On Apr 14, 2000 this sequence version replaced gi:7567976.
Contact: Dixon RA
Plant Biology Division
The Samuel Roberts Noble Foundation
2510 Sam Noble Parkway, Ardmore, OK 73402, USA
Tel: 580 221 7302
Fax: 580 221 7380
Email: radixon@noble.org
Insert Length: 752 Std Error: 0.00
Plate: 062 row: A column: 03
Seq primer: TCACACAGGAAACAGCTATGAC.
FEATURES Location/Qualifiers
source 1. .52
/organism="Medicago truncatula"

```
                     /db_xref="taxon:3880"
                     /clone="NF062A03ST"
                     /clone_lib="Developing stem"
                     /tissue_type="stem"
                     /dev_stage="Pooled developmental"
                     /note="Vector: Lambda Zap; Contains a mixture of
                     internodal stem segments"
BASE COUNT       12 a      22 c       2 g      16 t
ORIGIN


  Query Match             68.7%;  Score 15.8;  DB 10;  Length 52;
  Best Local Similarity   89.5%;  Pred. No. 4.5e+03;
  Matches   17;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       4 aacacccgctctctcgcaa 22
           ||||||| ||||||| |||
Db       9 AACACCCACTCTCTCACAA 27


RESULT   2
AU107235
LOCUS       AU107235       50 bp    mRNA            EST       05-APR-2001
DEFINITION  AU107235 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
            LNG07514, mRNA sequence.
ACCESSION   AU107235
VERSION     AU107235.1  GI:13556756
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 50)
  AUTHORS   Suzuki,Y., Tsunoda,T., Taira,H., Mizushima-Sugano,J., Sese,J., Hata
            ,H., Ota,T., Isogai,T., Tanaka,T., Nakamura,Y., Morishita,S., Okubo
            ,K., Suyama,A. and Sugano,S.
  TITLE     Fine Structural analysis of transcription start sites of human
            mRNAs using full-length enriched and 5'-end enriched cDNA libraries
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp
            Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano
            ,S. Construction and characterization of a full length-enriched and
            a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).
FEATURES             Location/Qualifiers
     source          1. .50
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="LNG07514"
                     /clone_lib="Sugano Homo sapiens cDNA library"
BASE COUNT       16 a      15 c       9 g      10 t
ORIGIN


  Query Match             63.5%;  Score 14.6;  DB 10;  Length 50;
  Best Local Similarity   81.0%;  Pred. No. 1.4e+04;
  Matches   17;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       2 agaacacccgctctctcgcaa 22
           |||||| || ||||||| | |
Db      22 AGAACAACCTCTCTCTCCA 42


RESULT   3
AI859778
LOCUS       AI859778       40 bp    mRNA            EST       07-MAR-2000
DEFINITION  wm21e09.x1 NCI_CGAP_Ut4 Homo sapiens cDNA clone IMAGE:2436616 3'
            similar to TR:Q08380 Q08380 MAC-2 BINDING PROTEIN PRECURSOR. ;,
            mRNA sequence.
ACCESSION   AI859778
VERSION     AI859778.1  GI:5513479
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 40)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html

            Trace considered overall poor quality
            Insert Length: 1070   Std Error: 0.00
            Seq primer: -40UP from Gibco
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .40
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2436616"
                     /clone_lib="NCI_CGAP_Ut4"
                     /tissue_type="serous papillary carcinoma, high grade, 2
                     pooled tumors"
                     /lab_host="DH10B"
                     /note="Organ: uterus; Vector: pCMV-SPORT6; Site_1: SalI;
                     Site_2: NotI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 1.48 kb. Life Technologies catalog #:
                     11542-016"
BASE COUNT       13 a      13 c       9 g       5 t
ORIGIN


  Query Match             60.9%;  Score 14;  DB 10;  Length 40;
  Best Local Similarity   77.3%;  Pred. No. 2.6e+04;
  Matches   17;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       2 agaacacccgctctctcgcaaa 23
           ||||||||| || || |  |||
Db       7 AGAACACCCTCTGTCCCTGAAA 28


RESULT   4
AZ463268
LOCUS       AZ463268       40 bp    DNA             GSS       04-OCT-2000
DEFINITION  1M0272B01F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC1M0272B01 F, DNA sequence.
ACCESSION   AZ463268
VERSION     AZ463268.1  GI:10621393
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 40)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
            and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
```

University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0272 row: B column: 01
Seq primer: CGTTGTAAAACGACGGCCAGT
Class: plasmid ends
High quality sequence stop: 40.

FEATURES Location/Qualifiers
source 1. .40
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC1M0272B01"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptored DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptored mouse DNA was annealed to adaptored vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 5 a 20 c 2 g 13 t
ORIGIN

Query Match 59.1%; Score 13.6; DB 13; Length 40;
Best Local Similarity 80.0%; Pred. No. 3.8e+04;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```
Qy      1 gagaacacccgctctctcgc 20
          || | ||||| ||||||| |
Db      2 GAAAGCACCCTCTCTCTCTC 21
```

RESULT 5
AQ025206/c
LOCUS AQ025206 34 bp DNA GSS 23-AUG-2000
DEFINITION EP(3)1249 Drosophila melanogaster EP line Drosophila melanogaster genomic Sequence recovered from 5' end of P element, DNA sequence.
ACCESSION AQ025206
VERSION AQ025206.1 GI:3265558
KEYWORDS GSS.
SOURCE fruit fly.
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 34)
AUTHORS Liao,G.-C., Rehm,E.J. and Rubin,G.M.
TITLE Insertion site preferences of the P transposable element in Drosophila melanogaster
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3347-3351 (2000)
MEDLINE 20202638
COMMENT Contact: Gerald Rubin
Berkeley Drosophila Genome Project
University of California, Berkeley
LSA Building, Berkeley, CA 94720-3200, USA
Fax: 5106439947
Email: gerry@fruitfly.berkeley.edu
Sequence recovery method was inverse PCR.

Sequence orientation is forward strand relative to 5' end of P element

The P element insertion position is base 27 in the 34 bases. This insertion position refers to the first base of the 8 base target recognition sequence.
Class: transposon-tagged.

FEATURES Location/Qualifiers
source 1. .34
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone_lib="Drosophila melanogaster EP line"
/note="Inverse PCR was performed on Drosophila melanogaster strains each of which contains a single EP transposable element insertion. (The generation of these insertion strains is described in Rorth P, Szabo K, Bailey A, Laverty T, Rehm J, Rubin GM, Weigmann K, Milan M, Benes V, Ansorge W, Cohen SM. 1998. Systematic gain-of-function genetics in Drosophila. Development 6:1049-1057.) The resultant fragment for each strain was directly sequenced to determine the genomic sequence at the site of insertion. Details of the protocols used can be found at http://fruitfly.berkeley.edu/p_disrupt/inverse_pcr.html."

BASE COUNT 4 a 5 c 17 g 8 t
ORIGIN

Query Match 58.3%; Score 13.4; DB 13; Length 34;
Best Local Similarity 93.3%; Pred. No. 4.5e+04;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy      4 aacacccgctctctc 18
          | ||||||||||||||
Db     28 ACCACCCGCTCTCTC 14
```

RESULT 6
AZ514447
LOCUS AZ514447 49 bp DNA GSS 05-OCT-2000
DEFINITION 1M0361M14F Mouse 10kb plasmid UUGC1M library Mus musculus genomic clone UUGC1M0361M14 F, DNA sequence.
ACCESSION AZ514447
VERSION AZ514447.1 GI:10695859
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 49)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D.,Weiss,R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0361 row: M column: 14

```
            Seq primer: CGTTGTAAAACGACGGCCAGT
            Class: plasmid ends
            High quality sequence stop: 49.
FEATURES              Location/Qualifiers
     source           1. .49
                      /organism="Mus musculus"
                      /strain="C57BL/6J"
                      /db_xref="taxon:10090"
                      /clone="UUGC1M0361M14"
                      /clone_lib="Mouse 10kb plasmid UUGC1M library"
                      /sex="Male"
                      /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                      /note="Vector: PWD42nv; Purified genomic DNA from M.
                      musculus C57BL/6J (male) was obtained  from the Jackson
                      Laboratory Mouse DNA Resource
                      (http://www.jax.org/resources/documents/dnares/). The DNA
                      was hydrodynamically sheared by repeated passage through a
                      0.005 inch orifice at constant velocity. The sheared DNA
                      was blunt end-repaired with T4 DNA polymerase and T4
                      polynucleotide kinase. Adaptor oligonucleotides were
                      ligated to the blunt ends in high  molar excess. The
                      adaptored DNA was purified and size-selected for a 9.5 to
                      10.5 kb range using preparative agarose gel
                      electrophoresis. Vector DNA was prepared from a derivative
                      of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                      inducible derivative of plasmid R1. The vector was ligated
                      with adaptors complementary to the insert adaptors and
                      purified. The sheared, adaptored mouse DNA was annealed to
                      adaptored vector DNA, and transformed into
                      chemically-competent E. coli XL10-Gold (Stratagene) cells
                      and selected for ampicillin resistance."
BASE COUNT       15 a      22 c       8 g       4 t
ORIGIN


  Query Match             58.3%;  Score 13.4;  DB 13;  Length 49;
  Best Local Similarity   73.9%;  Pred. No. 4.7e+04;
  Matches   17;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy       1 gagaacacccgctctctcgcaaa 23
           | | |||| ||| ||||| || |
Db      19 GCGCACACACGCGCTCTCACACA 41


RESULT   7
AZ494986/c
LOCUS       AZ494986       60 bp    DNA              GSS       05-OCT-2000
DEFINITION  1M0330A09R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC1M0330A09 R, DNA sequence.
ACCESSION   AZ494986
VERSION     AZ494986.1  GI:10670081
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 60)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
            and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0330  row: A  column: 09
            Seq primer: CACACAGGAAACAGCTATGACC
            Class: plasmid ends
            High quality sequence stop: 60.
FEATURES              Location/Qualifiers
     source           1. .60
                      /organism="Mus musculus"
                      /strain="C57BL/6J"
                      /db_xref="taxon:10090"
                      /clone="UUGC1M0330A09"
                      /clone_lib="Mouse 10kb plasmid UUGC1M library"
                      /sex="Male"
                      /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                      /note="Vector: PWD42nv; Purified genomic DNA from M.
                      musculus C57BL/6J (male) was obtained  from the Jackson
                      Laboratory Mouse DNA Resource
                      (http://www.jax.org/resources/documents/dnares/). The DNA
                      was hydrodynamically sheared by repeated passage through a
                      0.005 inch orifice at constant velocity. The sheared DNA
                      was blunt end-repaired with T4 DNA polymerase and T4
                      polynucleotide kinase. Adaptor oligonucleotides were
                      ligated to the blunt ends in high  molar excess. The
                      adaptored DNA was purified and size-selected for a 9.5 to
                      10.5 kb range using preparative agarose gel
                      electrophoresis. Vector DNA was prepared from a derivative
                      of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                      inducible derivative of plasmid R1. The vector was ligated
                      with adaptors complementary to the insert adaptors and
                      purified. The sheared, adaptored mouse DNA was annealed to
                      adaptored vector DNA, and transformed into
                      chemically-competent E. coli XL10-Gold (Stratagene) cells
                      and selected for ampicillin resistance."
BASE COUNT       18 a       9 c      21 g      12 t
ORIGIN


  Query Match             57.4%;  Score 13.2;  DB 13;  Length 60;
  Best Local Similarity   83.3%;  Pred. No. 5.8e+04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 gagaacacccgctctctc 18
           ||||||  || ||| ||||
Db      19 GAGAACTCCAGCTTTCTC 2


RESULT   8
AZ303786/c
LOCUS       AZ303786       25 bp    DNA              GSS       29-SEP-2000
DEFINITION  1M0003G03R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC1M0003G03 R, DNA sequence.
ACCESSION   AZ303786
VERSION     AZ303786.1  GI:10339164
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 25)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
            and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
```

```
            Insert Length: 10000   Std Error: 0.00
            Plate: 0003  row: G  column: 03
            Seq primer: CACACAGGAAACAGCTATGACC
            Class: plasmid ends
            High quality sequence stop: 25.
FEATURES              Location/Qualifiers
     source           1. .25
                      /organism="Mus musculus"
                      /strain="C57BL/6J"
                      /db_xref="taxon:10090"
                      /clone="UUGC1M0003G03"
                      /clone_lib="Mouse 10kb plasmid UUGC1M library"
                      /sex="Male"
                      /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                      /note="Vector: PWD42nv; Purified genomic DNA from M.
                      musculus C57BL/6J (male) was obtained  from the Jackson
                      Laboratory Mouse DNA Resource
                      (http://www.jax.org/resources/documents/dnares/). The DNA
                      was hydrodynamically sheared by repeated passage through a
                      0.005 inch orifice at constant velocity. The sheared DNA
                      was blunt end-repaired with T4 DNA polymerase and T4
                      polynucleotide kinase. Adaptor oligonucleotides were
                      ligated to the blunt ends in high  molar excess. The
                      adaptored DNA was purified and size-selected for a 9.5 to
                      10.5 kb range using preparative agarose gel
                      electrophoresis. Vector DNA was prepared from a derivative
                      of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                      inducible derivative of plasmid R1. The vector was ligated
                      with adaptors complementary to the insert adaptors and
                      purified. The sheared, adaptored mouse DNA was annealed to
                      adaptored vector DNA, and transformed into
                      chemically-competent E. coli XL10-Gold (Stratagene) cells
                      and selected for ampicillin resistance."
BASE COUNT         2 a      2 c      9 g     12 t
ORIGIN


  Query Match             56.5%;  Score 13;  DB 13;  Length 25;
  Best Local Similarity   76.2%;  Pred. No. 6.5e+04;
  Matches   16;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       2 agaacacccgctctctcgcaa 22
           | ||||| | | ||| |||||
Db      22 AAAACACACACACTCACGCAA 2


RESULT   9
AA098077/c
LOCUS       AA098077        59 bp    mRNA            EST       25-OCT-1996
DEFINITION  mk18g04.r1 Soares mouse p3NMF19.5 Mus musculus cDNA clone
            IMAGE:493302 5', mRNA sequence.
ACCESSION   AA098077
VERSION     AA098077.1  GI:1643631
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 59)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT     Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:296750
            Seq primer: -28M13 rev2 from Amersham
            High quality sequence stop: 53.
FEATURES              Location/Qualifiers
     source           1. .59
                      /organism="Mus musculus"
                      /db_xref="taxon:10090"
                      /clone="IMAGE:493302"
                      /clone_lib="Soares mouse p3NMF19.5"
                      /dev_stage="19.5 dpc total fetus"
                      /lab_host="DH10B (ampicillin resistant)"
                      /note="Vector: pT7T3D (Pharmacia) with a modified
                      polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                      was primed with a Not I - oligo(dT) primer [5'
                      TGTTACCAATCTGAAGTGGGAGCGGCCGCATTTTTTTTTTTTTTTTTT 3'],
                      double-stranded cDNA was size selected, ligated to Eco RI
                      adapters (Pharmacia), digested with Not I and cloned into
                      the Not I and Eco RI sites of a modified pT7T3 vector
                      (Pharmacia). Library went through one round of
                      normalization to a Cot = 5. Library constructed by Bento
                      Soares and M.Fatima Bonaldo. RNA was kindly provided by
                      Dr. Minoru Ko (Wayne State University)."
BASE COUNT         9 a      12 c     21 g     17 t
ORIGIN


  Query Match             56.5%;  Score 13;  DB 10;  Length 59;
  Best Local Similarity   76.2%;  Pred. No. 7.1e+04;
  Matches   16;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       2 agaacacccgctctctcgcaa 22
           ||| ||| ||| | |||||
Db      59 AGAGCACACGCGCCCTCGCCA 39


RESULT  10
AA894695/c
LOCUS       AA894695        55 bp    mRNA            EST       06-APR-1998
DEFINITION  oj27d05.s1 NCI_CGAP_Kid3 Homo sapiens cDNA clone IMAGE:1493385 3'
            similar to TR:Q92510 Q92510 MYELOBLAST KIAA0236. ;contains TAR1.t3
            MSR1 repetitive element ;, mRNA sequence.
ACCESSION   AA894695
VERSION     AA894695.1  GI:3031096
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 55)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html

            Trace considered overall poor quality
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 1.
FEATURES              Location/Qualifiers
     source           1. .55
                      /organism="Homo sapiens"
                      /db_xref="taxon:9606"
```

/clone="IMAGE:1493385"
/clone_lib="NCI_CGAP_Kid3"
/lab_host="DH10B"
/note="Organ: kidney; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer, double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. mRNA source: 2 pooled kidneys. Library went through one round of normalization. Library constructed by Bento Soares and M. Fatima Bonaldo. "

BASE COUNT 9 a 6 c 30 g 9 t 1 others
ORIGIN

Query Match 55.7%; Score 12.8; DB 10; Length 55;
Best Local Similarity 87.5%; Pred. No. 8.6e+04;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 5 acacccgctctctcgc 20
||||||| ||| |||||
Db 22 ACACCCTCTCCCTCGC 7

RESULT 11
AZ342914
LOCUS AZ342914 26 bp DNA GSS 29-SEP-2000
DEFINITION 1M0076C22F Mouse 10kb plasmid UUGC1M library Mus musculus genomic clone UUGC1M0076C22 F, DNA sequence.
ACCESSION AZ342914
VERSION AZ342914.1 GI:10420628
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 26)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D.,Weiss,R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0076 row: C column: 22
Seq primer: CGTTGTAAAACGACGGCCAGT
Class: plasmid ends
High quality sequence stop: 26.
FEATURES Location/Qualifiers
source 1. .26
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC1M0076C22"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptored DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptored mouse DNA was annealed to adaptored vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 3 a 11 c 1 g 11 t
ORIGIN

Query Match 54.8%; Score 12.6; DB 13; Length 26;
Best Local Similarity 78.9%; Pred. No. 9.6e+04;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 2 agaacacccgctctctcgc 20
||||| | | ||||||| |
Db 1 AGAACTCTCTCTCTCTCTC 19

RESULT 12
TA194F01Q/c
LOCUS TA194F01Q 26 bp DNA GSS 13-DEC-2000
DEFINITION T. brucei sheared genomic DNA clone 194f01, reverse sequence, genomic survey sequence.
ACCESSION AL477302
VERSION AL477302.1 GI:11841328
KEYWORDS GSS.
SOURCE Trypanosoma brucei.
ORGANISM Trypanosoma brucei
Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
REFERENCE 1 (bases 1 to 26)
AUTHORS Hall,N., Bowman,S., Lennard,N.J., Doggett,J., Atkin,R., Chillingworth,C., Ormond,D., Harris,B., El-Sayed,N., Hou,L., Melville,S.E., Rajandream,M.A. and Barrell,B.G.
TITLE Direct Submission
JOURNAL Submitted (10-DEC-2000) Trypanosoma brucei genome sequencing project, Sanger Centre, The Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, E-mail: barrell@sanger.ac.uk and nh1@sanger.ac.uk
COMMENT Constructed at the Institute for Genomic Research (TIGR), Rockville, MD. Genomic DNA isolated from a cloned population of Trypanosoma brucei (TREU927/4 GUTat 10.1) was mechanically sheared to give a tight size distribution (
4 kb). The v + i method used for the library construction is described in detail in Smith, H. and Venter, J.C. (Making small insert libraries for whole genome shotgun sequencing projects. In Genome Sequencing: A Practical Approach, eds. M. Vaudin and B. Barrell, Oxford University Press, 1999).
Email: nelsayed@tigr.org
Details of T. brucei sequencing at the Sanger Centre are available at http://www.sanger.ac.uk/Projects/T_brucei/.
FEATURES Location/Qualifiers
source 1. .26
/organism="Trypanosoma brucei"
/strain="TREU927"
/db_xref="taxon:5691"
/clone="194f01"
BASE COUNT 0 a 3 c 11 g 12 t
ORIGIN

Query Match 54.8%; Score 12.6; DB 13; Length 26;
Best Local Similarity 78.9%; Pred. No. 9.6e+04;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```
Qy      5 acacccgctctctcgcaaa 23
          |||| ||| | | ||||||
Db     21 ACACGCGCACACACGCAAA 3


RESULT  13
AZ845779/c
LOCUS       AZ845779       35 bp    DNA             GSS       20-FEB-2001
DEFINITION  2M0145B13R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC2M0145B13 R, DNA sequence.
ACCESSION   AZ845779
VERSION     AZ845779.1  GI:13015687
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 35)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
            and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0145  row: B  column: 13
            Seq primer: CACACAGGAAACAGCTATGACC
            Class: plasmid ends
            High quality sequence stop: 35.
FEATURES             Location/Qualifiers
     source          1. .35
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC2M0145B13"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
BASE COUNT        3 a      1 c     15 g     16 t
ORIGIN


  Query Match             54.8%;  Score 12.6;  DB 13;  Length 35;
  Best Local Similarity   78.9%;  Pred. No. 1e+05;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      5 acacccgctctctcgcaaa 23
          |||||||||| | | || |
Db     35 ACACCCGCTCACACACACA 17


RESULT  14
AU106768/c
LOCUS       AU106768       50 bp    mRNA            EST       05-APR-2001
DEFINITION  AU106768 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
            HEP11938, mRNA sequence.
ACCESSION   AU106768
VERSION     AU106768.1  GI:13556289
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 50)
  AUTHORS   Suzuki,Y., Tsunoda,T., Taira,H., Mizushima-Sugano,J., Sese,J., Hata
            ,H., Ota,T., Isogai,T., Tanaka,T., Nakamura,Y., Morishita,S., Okubo
            ,K., Suyama,A. and Sugano,S.
  TITLE     Fine Structural analysis of transcription start sites of human
            mRNAs using full-length enriched and 5'-end enriched cDNA libraries
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp
            Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano
            ,S. Construction and characterization of a full length-enriched and
            a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).
FEATURES             Location/Qualifiers
     source          1. .50
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="HEP11938"
                     /clone_lib="Sugano Homo sapiens cDNA library"
BASE COUNT       12 a     11 c     17 g     10 t
ORIGIN


  Query Match             54.8%;  Score 12.6;  DB 10;  Length 50;
  Best Local Similarity   78.9%;  Pred. No. 1e+05;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      2 agaacacccgctctctcgc 20
          || ||||| |||| ||| |
Db     40 AGCACACCGGCTCCCTCTC 22


RESULT  15
AZ462085
LOCUS       AZ462085       32 bp    DNA             GSS       04-OCT-2000
DEFINITION  1M0269P08F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC1M0269P08 F, DNA sequence.
ACCESSION   AZ462085
VERSION     AZ462085.1  GI:10620210
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 32)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
            and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
```

University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0269 row: P column: 08
Seq primer: CGTTGTAAAACGACGGCCAGT
Class: plasmid ends
High quality sequence stop: 32.

FEATURES Location/Qualifiers
source 1. .32
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC1M0269P08"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptored DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptored mouse DNA was annealed to adaptored vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 7 a 14 c 1 g 10 t
ORIGIN

Query Match 53.9%; Score 12.4; DB 13; Length 32;
Best Local Similarity 92.9%; Pred. No. 1.2e+05;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy       5 acacccgctctctc 18
           |||||| |||||||
Db       3 ACACCCACTCTCTC 16
```

Search completed: March 9, 2002, 00:09:25
Job time: 11041 sec